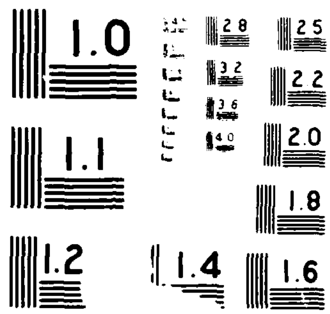
AD-A192 355 CONDITIONS FOR FINITE CONVERGENCE OF ALGORITHMS FOR NONLINEAR PROGRAMS AN.. (U) GEORGIA INST OF TECH ATLANTA PRODUCTION AND DISTRIBUTION RESE.. 1/1

UNCLASSIFIED F A AL-KHAYYAL ET AL. JAN 88 PDRC-88-01 F/G 12/4 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.2
1.4
1.6

(4)

AD-A192 355

DTIC FILE COPY

Georgia Institute of Technology

SEAL OF THE GEORGIA INSTITUTE OF TECHNOLOGY 1885

DTIC
ELECTE
MAR 1 1 1988
S D
D

PRODUCTION and DISTRIBUTION RESEARCH CENTER

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

SCHOOL OF INDUSTRIAL
AND SYSTEMS
ENGINEERING
GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM
OF GEORGIA
ATLANTA, GEORGIA 30332

88 2 19 086

PDRC Report Series 88-01
January 1988

CONDITIONS FOR FINITE CONVERGENCE OF ALGORITHMS FOR NONLINEAR PROGRAMS AND VARIATIONAL INEQUALITIES

by

*Faiz A. Al-Khayyal†
**Jerzy Kyparisis

PDRC 88-01

*School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332

**Florida International University

†Author's work was supported in part by the Office of Naval Research under Contract No. N00014-86-K-0173 and by general research development funds provided by the Georgia Institute of Technology.

CONDITIONS FOR FINITE CONVERGENCE OF ALGORITHMS FOR NONLINEAR PROGRAMS AND VARIATIONAL INEQUALITIES

Faiz A. Al-Khayyal†
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332

and

Jerzy Kyparisis
Department of Decision Sciences
and Information Systems
Florida International University
Miami, FL 33199

ABSTRACT

Algorithms for nonlinear programming and variational inequality problems are, in general, only guaranteed to converge in the limit to a Karush-Kuhn-Tucker point, in the case of nonlinear programs, or a solution in the case of variational inequalities. In this paper we derive sufficient conditions for nonlinear programs and variational inequalities such that any convergent algorithm can be modified to guarantee finite convergence to a solution. Our conditions are more general than existing results and, in addition, have wider applicability. Moreover, we note that our sufficient conditions are close to the related necessary conditions, and show by counterexamples that our main nondegeneracy assumptions cannot be relaxed.

Key Words: Convergence of algorithms, nonlinear programs, variational inequalities.

†Author's work was supported in part by the Office of Naval Research under Contract No. N00014-86-0173 and by general research development funds provided by the Georgia Institute of Technology.

CONDITIONS FOR FINITE CONVERGENCE OF ALGORITHMS FOR NONLINEAR PROGRAMS AND VARIATIONAL INEQUALITIES

Faiz A. Al-Khayyal
Georgia Institute of Technology

and

Jerzy Kyparisis
Florida International University

1. Introduction

Numerous algorithms have been devised for the solution of the optimization problem

$$\text{minimize } f(x) \text{ subject to } x \in S. \qquad (1.1)$$

Differences in the procedures generally exploit the assumed structure of the objective function f and the feasible set S. In this paper, we develop conditions on f, S, and the solution point under which <u>any</u> infinitely convergent algorithm can be modified to ensure finite convergence. Within the assumptions on f and S we show by counter-examples that the conditions on the solution point are tight. Our results are more general than earlier related results and have been extended to the problem of finding solutions to variational inequalities, which calls for finding a point $x^* \in S$ such that

$$G(x^*)^T (x-x^*) \geq 0 \quad \text{for all } x \in S, \qquad (1.2)$$

where G is a continuous mapping from R^n to R^n.

Problem (1.1) is called a nonlinear program whenever either f is a continuous nonlinear function or the set S is defined by a collection of continuous nonlinear inequalities and equations. Under a suitable constraint qualification (see, e.g., [3]), algorithms for nonlinear programming problems are, in general, only guaranteed to converge in the limit to a Karush-Kuhn-Tucker point (see, e.g., [20]), which may or may not be a local solution. For special cases, convergence to a local solution can be established (e.g., for convex programs [1]), or finite convergence can be assured (e.g., for conjugate direction algorithms applied to bounded quadratic programs using Armijo-type inexact line searches [20]). Algorithms for variational inequality problems typically are guaranteed to converge in the limit to a solution if such a solution is unique. Thus, in most cases, the existing algorithms converge to either Karush-Kuhn-Tucker or solution points in an *infinite number of* steps. We propose to modify these procedures by imbedding a subproblem (having linear objective function and constraints given by S) that should be solved every fixed number of steps. Under certain conditions, this suffices to guarantee finite convergence to a solution. This idea is reminiscent of the "spacer step" used in nonlinear programming to convert a locally convergent algorithm into a globally convergent one (see, e.g., Luenberger [19]).

Rates of convergence and conditions for finite convergence have been developed by a number of authors for gradient projection algorithms. Bertsekas [5] studied the Goldstein-Levitin-Polyak [16,18] gradient projection method for problem (1.1) with continuously differentiable objective and compact S defined by only upper and lower bounds on the

decision vector using Armijo type [2] inexact line searches. Bertsekas proved that if the sequence generated by the algorithm converges to a local minimizer which satisfies strict complementary slackness (SCS) and second order sufficiency conditions (SOSC), then the set of active constraints is identified in a finite number of steps. In [6], Bertsekas develops analogous results for the same problem when applying a projected Newton method. In a related paper, Gafni and Bertsekas [14] investigate convergence and rate of convergence of a generalization of the Goldstein-Levitin-Polyak (GLP) method, under the assumptions that f is smooth and real valued and that S is convex in a linear space, to allow different Hilbert norms to be used for the projection operation and the Frechet derivative operation. In a very thorough and rigorous study, Dunn [13] analyzes gradient projection algorithms of the GLP type when applied to a general convex set S in Hilbert space and a continuously differentiable real valued objective function f. When f has a unique minimizer ξ in S, then convergence rates can be obtained and are dependent on how the function

$$0 \le \nu(\sigma) = \inf\{f(x) - f(\xi) : x \in S, \|x - \xi\| \ge \sigma\} \qquad (1.3)$$

grows with increasing $\sigma > 0$. Note that the growth of $\nu(\sigma)$ depends on the structure of f, S, and the norm on the Hilbert space. Dunn proves that if for some $\nu > 0$ there exists a scalar $B > 0$ such that $\nu(\sigma) \ge B\sigma^{\nu}$ for all $\sigma > 0$, then the process $\{f(x_n)\}$ converges to $\inf\{f(x) : x \in S\}$ [illegible], superlinearly, or finitely if $\nu > 2$, $1 < \nu < 2$ or $\nu = 1$, respectively, where $\{x_n\}$ is the sequence of points generated by the GLP

algorithm. The same approach has been successfully used by Dunn to study convergence of other algorithms [10,11,12].

Conditions that use theoretical functions such as (1.3) are useful but difficult to apply. One of the simplest conditions was also provided by Dunn [11] who proved that the gradient projection method identifies the optimal active constraint set in a finite number of iterations whenever SCS holds at the solution. Burke and More [7] extend this result by showing that the optimal active constraints are identified if and only if the projected gradients converge to zero. Finally, Calamai and More [8] generalize the results of Bertsekas [5] by "showing that if the projected gradients converge to zero and if the iterates converge to a nondegenerate point, then the optimal active and binding constraints of a general linearly constrained problem are identified in a finite number of iterations. This result is independent of the method used to generate the iterates and can be applied to other linearly constrained problems." Their proof does not require SOSC and permits S to be a general polyhedral set. By invoking the gradient projection method on selected iterations, Calamai and More were able to develop a finite procedure for determining the global solution of a general quadratic programming problem.

In this paper, we generalize the results of Calamai and More [8] to problems with nonpolyhedral feasible sets and to algorithms which are not based on projected gradient calculations. Moreover, the nature of our approach allows us to extend our results to algorithms for solving variational inequalities. While many of the ideas contained in our results are not surprising, we have not seen them presented elsewhere and

attempt here to structure and characterize general yet useful results for R^n. In addition, our results have much simpler proofs than related results in the literature [5,8,11]. The remainder of this paper is organized as follows. Section 2 is concerned with conditions for finite convergence of algorithms for nonlinear programs with linear constraints. The case of nonlinear programs with nonlinear inequality constraints is dealt with in Section 3, and Section 4 extends the results in the preceding sections to variational inequality problems.

2. NONLINEAR PROGRAMS WITH LINEAR CONSTRAINTS

Consider the linearly constrained nonlinear programming problem of finding a point $x \in R^n$ that will locally

$$\text{minimize } f(x) \quad \text{subject to } Ax = a, \quad Bx \geq b \qquad (P)$$

where $f: R^n \to R^1$ is continuously differentiable, A and B are real $p \times n$ and $m \times n$ matrices, and a and b are real $p \times 1$ and $m \times 1$ vectors, respectively. Let $C = \{x: Ax=a, Bx\geq b\}$ be the feasible set of P.

In general, we seek a <u>local solution</u> to the problem; that is, a point x^* for which there is some neighborhood N (of x^*) such that $f(x) \geq f(x^*)$ for all x in $C \cap N$. If equality holds only for $x = x^*$, then x^* is called a <u>strict local solution</u> of problem P. In contrast, x^* is called an <u>isolated local solution</u> (or a <u>locally unique solution</u>) of problem P if it is the only local solution of P in some neighborhood of x^*. The distinction between strict and isolated local solutions for general, nonlinearly constrained problems has been investigated by Robinson [22] and by Kyparisis and Fiacco [17] who developed sufficient conditions for

each type of solution. As might be expected, the distinction disappears for special cases; e.g., when f is a quadratic function in P (see [17]).

In the sequel, we review several definitions and results from convex analysis and optimization theory which are needed for the development of our conditions. The reader may consult, e.g., Bazaraa and Shetty [3] for complete proofs and further details.

It is well-known [3] that if x^* is a local solution of problem P, then there exists a (dual) vector $(u,v) \in R^{p+m}$ such that (x^*,u,v) satisfies the following Karush-Kuhn-Tucker (KKT) conditions for P

(i) $\nabla f(x^*) = A^T u + B^T v, \quad v \geq 0,$

(ii) $Ax^* = a, \quad Bx^* \geq b,$

(iii) $v^T(Bx^* - b) = 0.$

Denote the index set of binding inequality constraints at x^* by $I^* = \{i: B_i x^* = b_i\}$, where B_i denotes the ith row of matrix B. The corresponding submatrix of [B,b] is denoted by $[B^*,b^*]$. The <u>cone of feasible directions</u> of C at x^* is given by

$$F(x^*) = \{z: \quad z \neq 0, \ Az = 0, \ B^* z \geq 0\}$$

and the <u>polar cone</u> of $F(x^*)$ is given by

$$F^*(x^*) = \{y: y = A^T u + B^{*T} v^*, \ u \in R^p, \ v^* \geq 0\},$$

where v^* denotes the subvector of v with indices in I^*. Notice that $F^*(x^*)$ is simply the cone spanned by the negative gradients of the

binding constraints at x^*. Now, if x^* is a local solution, then the KKT conditions state that we must have

$$-\nabla f(x^*) \in F^*(x^*), \quad x^* \in C. \tag{2.1}$$

Later in this section we utilize the condition that $-\nabla f(x^*)$ is in the interior of $F^*(x^*)$, denoted by $\operatorname{int} F^*(x^*)$. The next proposition states the necessary and sufficient conditions for $\operatorname{int} F^*(x^*)$ to be nonempty.

PROPOSITION 1. The following conditions are equivalent:

(a) $\operatorname{int} F^*(x^*) \neq \emptyset$

(b) x^* is an extreme point of C;

(c) $\operatorname{rank}\,[A^T, B^{*T}] = n$; i.e., there are n linearly independent binding constraints at x^*.

One can verify that the interior of $F^*(x^*)$ is given by

$$\operatorname{int} F^*(x^*) = \{-y\colon\ y = A^T u + B^{*T} v^*, \text{ where } v^* \geq 0 \text{ is such that } v^{*T} B^* z > 0 \quad \text{for all } z \in F(x^*)\}.$$

If, in addition, the matrix $[A^T, B^{*T}]$ is $n \times n$, i.e., all the binding constraints are linearly independent (in view of (c) above), then

$$\operatorname{int} F^*(x^*) = \{-y\colon\ y = A^T u + B^{*T} v^*,\ v^* > 0\} \tag{2.2}$$

In general, however, int $F^*(x^*)$ only contains the set [illegible] known theorems of the alternative do not seem to be useful in representing int $F^*(x^*)$ in a simpler form for the general case.

The next result provides several characterizations for the condition $-\nabla f(x^*) \in \text{int } F^*(x^*)$ used in our main theorem.

PROPOSITION 2. Consider the conditions

(a) $-\nabla f(x^*) \in \text{int } F^*(x^*)$;

(b) $\nabla f(x^*)^T z > 0$ for all $z \in F(x^*)$ and x^* is an extreme point of C;

(c) x^* is the unique solution of $\min\{\nabla f(x^*)^T x : x \in C\}$;

(d) x^* is an isolated local solution of $\min\{f(x) : x \in C\}$.

Then, (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) $\Rightarrow$ (d).

Proof. (a) $\Rightarrow$ (b). Condition (a) implies that x^* is an extreme point of C by Proposition 1 and that $\nabla f(x^*)^T z > 0$ for all $z \in F(x^*)$ by definition of int $F^*(x^*)$.

(b) $\Rightarrow$ (c). Condition (b) implies $\nabla f(x^*)^T(x - x^*) > 0$ for all $x \in C$, $x \neq x^*$, i.e., $\nabla f(x^*)^T x > \nabla f(x^*)^T x^*$ for all $x \in C$, $x \neq x^*$.

(c) $\Rightarrow$ (a). Condition (c) implies $\nabla f(x^*)^T(x - x^*) > 0$ for all $x \in C$, $x \neq x^*$, i.e., $\nabla f(x^*)^T z > 0$ for all $z \in F(x^*)$. By definition of int $F^*(x^*)$, this means that $-\nabla f(x^*) \in \text{int } F^*(x^*)$.

Finally, (a) implies that the KKT conditions hold at x^*, and (b) implies that $\{z : \nabla f(x^*)^T z \leq 0,\ z \in F(x^*)\} = \emptyset$, so that, by Corollary 1.10 in [17], x^* is an isolated local solution of $\min\{f(x) : x \in C\}$ and thus (a), (b), or (c) imply (d).

[illegible] The converse follows immediately from the [illegible]

[illegible] the assumption that x^* is an extreme point of C, [illegible] $\nabla f(x^*)$ to be only in the relative interior of $F^*(x^*)$ [illegible] counter examples whereby $x^k \to x^*$ and f is such [illegible] $\nabla f(x^k)$ is linearly independent of the gradients of [illegible] constraints at a nonextreme point x^*. To illustrate, suppose [illegible] sequence $\{x^k\}$ and f satisfy

$$x^{k+1} = x^k + \frac{1}{2} \|x^* - x^k\| (x^* - x^k)$$

$$\nabla f(x^k) = \nabla f(x^*) + \frac{1}{k} e$$

where $e = (1,1,\ldots,1)^T \in R^n$ and $\|\cdot\|$ denotes any norm defined on R^n. If x^* is not an extreme point, it follows that x^* does not solve (2.3) for any finite k. We note that the nondegeneracy assumption of Calamai and More [8] is equivalent to assuming that $F^*(x^*)$ has a nonempty interior and that $-\nabla f(x^*) \in \text{int } F^*(x^*)$.

The preceding theorem suggests that any algorithm which generates a sequence of points converging to an extreme point x^* of a polyhedral set can be modified to guarantee finite convergence, when $-\nabla f(x^*) \in \text{int } F^*(x^*)$, by simply adding the subproblem of solving the linear program (2.3) every, say, n steps. We are guaranteed to reach the threshold K in a finite number of iterations and this is followed by a single linear program to determine x^*.

3. NONLINEAR PROGRAMS WITH NONLINEAR INEQUALITY CONSTRAINTS

Consider the nonlinear programming problem with nonlinear inequality constraints and linear equality constraints

$$\text{minimize } f(x) \text{ subject to } Ax = a,\ g_i(x) \geq 0,\ i=1,\ldots,m. \qquad \text{(NLP)}$$

where $f:R^n \to R^1$ and $g_i:R^n \to R^1$ are continuously differentiable, and, for simplicity, it is assumed that all the constraints are binding at x^*. Let $X = \{x: Ax = a,\ g_i(x) \geq 0,\ i=1,\ldots,m\}$ be the feasible set of NLP.

It is well known [3] that if x^* is a local solution of problem NLP and an appropriate constraint qualification holds then there exists a vector $(u,v) \in R^{p+m}$ such that (x^*,u,v) satisfies the following Karush-Kuhn-Tucker (KKT) conditions for NLP:

$$\text{(i)} \quad \nabla f(x^*) = A^T u + \nabla g(x^*)^T v, \quad v \geq 0,$$

$$\text{(ii)} \quad Ax^* = a, \quad g(x^*) \leq 0,$$

$$\text{(iii)} \quad v^T g(x^*) = 0,$$

where ∇g denotes the Jacobian matrix of the mapping $g = (g_1, g_2, \ldots, g_m)^T$.

In our results for this more general nonlinear program, we shall use the cone of tangents of X at x^*, denoted by $T(x^*)$, instead of the cone of feasible directions $F(x^*)$. When the constraints are linear, $T(x^*) = F(x^*)$, but in general only $F(x^*) \subset T(x^*)$ holds. The following proposition is well known.

PROPOSITION 3. Assume that the Kuhn Tucker Constraint Qualification (KTCQ) [3] holds at x^* for NLP. Then,

$$T(x^*) = \{z \mid z \neq 0, \; Az = 0, \; \nabla g_i(x^*)^T z \leq 0, \; i=1, \ldots, m\}$$

and the polar cone of $T(x^*)$ is given by

$$T^*(x^*) = \{y \mid y = A^T u + \nabla g(x^*)^T v, \; u \in R^p, \; v \geq 0\}$$

Of course, a number of other constraint qualifications which imply KTCQ can be used, for example, the Linear Independence CQ or, in the case of convex constraints g_i, the Slater CQ (see [3]). Note also that if x^* is a local solution and KTCQ holds, then the KKT conditions imply that

$$-\nabla f(x^*) \in T^*(x^*), \quad x^* \in X. \tag{3.1}$$

PROPOSITION 4. Assume that the KTCQ holds at x^*. Then, the following conditions are equivalent:

(a) $\text{int } T^*(x^*) \neq \emptyset$;

(b) $\text{rank } [A^T, \nabla g(x^*)^T] = n$; i.e., there are n linearly independent binding constraints at x^*.

The formula for $\text{int } T^*(x^*)$ is identical to that for $\text{int } F^*(x^*)$ with $\nabla g(x^*)$ substituted for B^*. Therefore, the discussion in Section 2 of various representations of $\text{int } F^*(x^*)$ applies here to $T^*(x^*)$.

The following lemma characterizes the key condition $-\nabla f(x^*) \in \text{int } T^*(x^*)$ that will be used later.

PROPOSITION 5. Assume that the KTCQ holds at x^* and consider the following conditions:

(a) $-\nabla f(x^*) \in \text{int } T^*(x^*)$;

(b) $\nabla f(x^*)^T z > 0$ for all $z \in T(x^*)$;

(c) x^* is the unique solution of $\min\{\nabla f(x^*)^T x : x \in X\}$;

(d) x^* is an isolated local solution of $\min\{f(x) : x \in X\}$.

Then, (a) $\Leftrightarrow$ (b) $\Rightarrow$ (c) and (d).

<u>Proof</u> (a) $\Leftrightarrow$ (b). Follows from the definition of $\text{int } T^*(x^*)$.

(b) $\Rightarrow$ (c). Condition (b) implies $\nabla f(x^*)^T(x - x^*) > 0$ for all $x \in x^* + T(x^*)$, which in turn implies (c) since $X \subset x^* + T(x^*)$.

(a) $\Rightarrow$ (d). Condition (a) implies that the KKT conditions hold at x^* and

that $\{z: \nabla f(x^*)^T z \le 0,\ z \in T(x^*)\} = \emptyset$, so that, by Corollary 4.10 in [17], x^* is an isolated local solution of $\min\{f(x): x \in X\}$.

The main result of this section parallels Theorem 1 and gives conditions under which an infinitely convergent algorithm for NLP can be modified to converge in a finite number of steps.

THEOREM 2. Assume that f and g are continuously differentiable and that the KTCQ holds at $x^* \in X$. Suppose that $\{x^k\} \subset X$, $x^k \to x^*$, and rank $[A^T, \nabla g(x^*)^T] = n$. If $-\nabla f(x^*) \in \text{int } T^*(x^*)$, then there is a positive integer K such that, for all $k > K$, x^* uniquely solves the nonlinear program

$$\text{minimize} \quad \nabla f(x^k)^T x \quad \text{subject to} \quad x \in X. \tag{3.2}$$

Conversely, if x^* solves (3.2) for all k larger than some number K, then x^* is a Karush-Kuhn-Tucker point of problem NLP.

Proof: By the continuous differentiability of f, we have $-\nabla f(x^k) \in \text{int } T^*(x^*)$ for sufficiently large k, since $x^k \to x^*$ and $-\nabla f(x^*) \in \text{int } T^*(x^*)$. Thus, $\nabla f(x^k)^T z > 0$ for all $z \in T(x^*)$; i.e., $\nabla f(x^k)^T (x-x^*) > 0$ for all $x \in x^* + T(x^*)$. Since $X \subset x^* + T(x^*)$, then $\nabla f(x^k)^T x > \nabla f(x^k)^T x^*$ for all $x \in X$, $x \neq x^*$. The converse follows immediately from continuous differentiability of f.

Remark. Analogous to Theorem 1, in the above theorem, we cannot relax the assumption that there are n linearly independent binding constraints at x^*, even if we require $-\nabla f(x^*)$ to be only in the relative interior of $T^*(x^*)$.

Theorem 2 suggests that any algorithm for NLP which generates a sequence of points converging to a point x^* can be modified to guarantee finite convergence, when $-\nabla f(x^*) \in \text{int } T^*(x^*)$, by adding the nonlinear program (3.2) with linear objective function every, say, n steps. This is guaranteed to yield a local solution x* in a finite number of iterations.

4. EXTENSIONS TO VARIATIONAL INEQUALITIES

In this section we consider the variational inequality problem of finding a point $x^* \in S$ such that

$$G(x^*)^T(x - x^*) \geq 0 \qquad \text{for all } x \in S, \tag{VI}$$

where $G: R^n \to R^n$ is continuous, S is a convex feasible set and x^* is called a solution of VI. Note that x^* solves the following convex program

$$\text{minimize} \quad G(x^*)^T x \quad \text{subject to} \quad x \in S. \tag{CP}$$

Furthermore, x^* is called an isolated solution (or a locally unique solution) of problem VI if it is the only solution of VI in some neighborhood of x^*.

Problem VI is, in a sense, a generalization of the nonlinear program NLP (and also of P). For example, if x^* is a local solution of NLP and if feasible set X in NLP is convex, then x^* solves VI with $G(x) = \nabla f(x)$ and $S = X$. Indeed, if the objective function f and set X in NLP are

convex, then VI subsumes NLP as the special case where $G(x) = \nabla f(x)$. On the other hand, problem VI directly generalizes the nonlinear (and therefore also linear) complementarity problem, which is obtained by setting $S = \{x: x \geq 0\}$. Algorithms for variational inequalities in most cases converge in the limit to a solution of VI if such a solution is unique (see Pang and Chan [21] and Dafermos [9]). Thus, finite convergence of these algorithms is also of interest in the VI case.

The results from Sections 2 and 3 directly extend in most cases to problem VI with an appropriate feasible set. Consider problem VI with a polyhedral feasible set $S = C$ (as defined in Section 2). Then CP becomes a linear program with optimal solution x^*. Thus, there exists a vector $(u,v) \in R^{p+m}$ such that (x^*,u,v) satisfies the following generalized Karush-Kuhn-Tucker (GKKT) conditions for VI:

(i) $G(x^*) = A^T u + B^T v, \quad v \geq 0,$

(ii) $Ax^* = a, \quad Bx^* \leq b,$

(iii) $v^T(Bx^* - b) = 0.$

The cones $T(x^*)$ and $T^*(x^*)$ are defined exactly as in Section 2, and Proposition 1 and comments following it apply without any changes. The GKKT conditions may be expressed in the form

$$G(x^*) \in T^*(x^*), \quad x^* \in C. \tag{4.1}$$

A strengthened condition $G(x^*) \in \operatorname{int} T^*(x^*)$ is characterized in the next result.

PROPOSITION 6. Consider the conditions:

(a) $-G(x^*) \in \text{int } F^*(x^*)$;

(b) $G(x^*)^T z > 0$ for all $z \in F(x^*)$ and x^* is an extreme point of C;

(c) x^* is the unique solution of $\min\{G(x^*)^T x : x \in C\}$;

(d) x^* is an isolated solution of $G(x)^T (z - x) \geq 0$ for all $z \in C$.

Then, (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) $\Rightarrow$ (d).

Proof. The first three equivalencies are proved exactly as in Proposition 2. The last implication follows from Theorem 2.3 in Tobin [23], since the GKKT conditions hold at x^* and since $\{z : G(x^*)^T z = 0, z \in F(x^*)\} = \{0\}$. ✓

The conditions under which an infinitely convergent algorithm can be modified to converge in a finite number of steps are stated in the theorem below, whose proof is analogous to the proof of Theorem 1 in Section 2 and is omitted.

THEOREM [illegible]. Assume that G is continuous and suppose that $\{x^k\} \subset C$, $x^k \to x^*$ and x^* is an extreme point of C. If $-G(x^*) \in \text{int } F^*(x^*)$, then there is a positive integer K such that, for all $k \geq K$, x^* uniquely solves the linear program

$$\text{minimize} \quad G(x^k)^T x \quad \text{subject to} \quad x \in C \qquad ([illegible])$$

Conversely, if x^* solves ([illegible]) for all k larger than some number K, then x^* is a generalized Karush-Kuhn-Tucker point of problem VI.

Remark. As in Theorem 2, the assumption that x^* is an extreme point of C cannot be relaxed even if we require $G(x^*)$ to be in the relative interior of $F^*(x^*)$.

Theorem 3 shows that any algorithm for VI with a polyhedral feasible set $S = C$ which generates a sequence of points converging to an extreme point x^* of C can be modified to guarantee finite convergence, provided that $G(x^*) \in \text{int } F^*(x^*)$. This is accomplished by solving an additional linear program (3.7) every, say, n steps (see also the comments following Theorem 2).

Consider now problem VI with a convex feasible set $S = X$, as defined in Section 3 with the additional assumption that each g_i is concave. For simplicity, we assume that all the constraints are binding at x^*. Thus, if the constraint qualification KTCQ holds at x^* (see Section 3), then there exists a vector $(u,v) \in R^{p+m}$ such that (x^*,u,v) satisfies the following generalized Karush-Kuhn-Tucker (GKKT) conditions for VI:

(i) $G(x^*) - A^T u + \nabla g(x^*)^T v = 0$,

(ii) $Ax^* = a$, $g(x^*) = 0$,

(iii) $v^T g(x^*) = 0$.

The definitions of cones $F(x^*)$ and $F^*(x^*)$ given in Section 3 as well as Propositions 3 and 4 and the comments following them in the same section apply without any changes. Also, the GKKT conditions may be expressed in the form

$$G(x^*) \in F^*(x^*), \quad x^* \in X \qquad (4.3)$$

The next result characterizes the key assumption in the subsequent main theorem, namely the condition $-G(x^*) \in \text{int } \Gamma^*(x^*)$. Its proof is similar to the proofs of Propositions 5 and 6.

PROPOSITION 7. Assume that the KTCQ holds at x^* and consider the following conditions:

(a) $-G(x^*) \in \text{int } \Gamma^*(x^*)$;

(b) $G(x^*)^T z > 0$ for all $z \in \Gamma(x^*)$;

(c) x^* is the unique solution of $\min\{G(x^*)^T x : x \in X\}$;

(d) x^* is an isolated solution of $G(x)^T(z - x) \geq 0$ for all $z \in X$.

Then, (a) $\Leftrightarrow$ (b) $\Rightarrow$ (c) and (d).

The conditions under which an infinitely convergent algorithm can be modified to converge in a finite number of steps are given below. The proof of this result is analogous to the proofs of Theorems 2 and 3. Observe also that since $G(x^*)^T x$ is linear and X is convex, in view of Proposition 7, part (c), x^* must be an extreme point of X. Recall that this was the case for problem VI with polyhedral feasible set C (see Theorem 3).

THEOREM 4. Assume that G is continuous, g is concave and that the KTCQ holds at x^*. Suppose that $\{x^k\} \subset X$, $x^k \to x^*$, and rank $[A^T, \nabla g(x^*)^T] = n$. If $-G(x^*) \in \text{int } \Gamma^*(x^*)$, then there is a positive integer K such that, for all $k \geq K$, x^* uniquely solves the convex program

$$\text{minimize } G(x^k)^T x \text{ subject to } x \in X \qquad (4.4)$$

conversely, if x^* solves (4.4) for all k larger than some number K, then x^* is a generalized Karush-Kuhn-Tucker point of problem IV.

Remark. As in Theorem 2, we cannot relax the assumption that there are n linearly independent binding constraints at x^*, even if we require $-G(x^*)$ to be only in the relative interior of $\Gamma^*(x^*)$.

Theorem 4 shows that any algorithm for VI with a convex set X which generates a sequence of points converging to x^* can be modified to guarantee finite convergence, when $-G(x^*) \in \text{int } \Gamma^*(x^*)$, by solving an additional convex program (4.4) with linear objective function every, say, n steps.

5. Concluding Remarks

We have derived necessary conditions and sufficient conditions under which any convergent algorithm for certain nonlinear programming problems and variational inequality problems can be modified to ensure finite convergence. Throughout we assumed the objective function to be continuously differentiable and considered generalizations of the constraints to extend our results. Using our approach, it appears that the most general constraints we can assume are linear equality and nonlinear inequality satisfying a constraint qualification. Moreover, we cannot relax the nondegeneracy assumptions on the solution points (namely, the polar cones are full dimensional and contain the negative gradients in their interior) and still ensure finiteness. In this sense, our results cannot be further weakened without additional assumptions.

Some existing algorithms may immediately benefit from results contained in this paper. For example, when the Al-Khayyal and Falk [2]

algorithm is applied to problems having global solutions that are nondegenerate extreme points, then the procedure is finite without modification. This is because the algorithm incorporates a step which solves a subproblem analogous to problem (2.3) before branching. Further refinements of the procedure and a specialized proof of finite convergence for nondegenerate linear complementarity problems have been investigated by Al Khayyal [1]

Acknowledgements

We are indebted to Garth McCormick, Hoang Tuy, and Hanif Sherali for helpful comments and suggestions.

REFERENCES

[1] Al-Khayyal, F. A., "Note on Solving Linear Complementarity Problems as Jointly Constrained Bilinear Programs," Working Paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Ga. 30332 (1987).

[2] Al-Khayyal, F. A. and J. E. Falk, "Jointly Constrained Biconvex Programming," Mathematics of Operations Research 8 (1983) 273-286.

[3] Bazaraa, M. S. and C. M. Shetty, Foundations of Optimization, Lecture Notes in Economics and Mathematical Systems 122, Springer-Verlag, Berlin, 1976.

[4] Bazaraa, M. S. and C. M. Shetty, Nonlinear Programming Theory and Algorithms, John Wiley & Sons, New York, 1979.

[5] Bertsekas, D. P., "On the Goldstein-Levitin-Polyak Gradient Projection Method," IEEE Transactions on Automatic Control 21 (1976) 174-184.

[6] Bertsekas, D. P., "Projected Newton Methods for Optimization Problems with Simple Constraints," SIAM Journal on Control and Optimization 20 (1982) 221-246.

[7] Burke, J. V. and J. J. Moré, "On the Identification of Active Constraints," Argonne National Laboratory, Mathematics and Computer Science Division Report ANL/MCS-TM-82, Argonne, Il, 1986.

[8] Calamai, P. H. and J. J. Moré, "Projected Gradient Methods for Linearly Constrained Problems," Mathematical Programming 39 (1987) 93-116.

[9] Dafermos, S., "An Iterative Scheme for Variational Inequalities," Mathematical Programming 26 (1983) 40-47.

[10] Dunn, J. C., "Rates of Convergence for Conditional Gradient Algorithms Near Singular and Nonsingular Extremals," SIAM Journal on Control and Optimization 17 (1979) 187-211.

[11] Dunn, J. C., "Convergence Rates for Conditional Gradient Sequences Generated by Implicit Step Length Rules," SIAM Journal on Control and Optimization 18 (1980) 473-487.

[12] Dunn, J. C., "Newton's Method and the Goldstein Step Length Rule for Constrained Minimization Problems," SIAM Journal on Control and Optimization 18 (1980) 659-674.

[13] Dunn, J. C., "Global and Asymptotic Convergence Rate Estimates for a Class of Projected Gradient Processes," SIAM Journal on Control and Optimization 19 (1981) 368-400.

[14] Dunn, J. C., "On the Convergence of Projected Gradient Processes to Singular Critical Points," Journal of Optimization Theory and Applications 55 (1987) 203-216.

[15] Gafni, E. M. and D. P. Bertsekas, "Two Metric Projection Methods for Constrained Optimization," SIAM Journal on Control and Optimization 22 (1984) 936-964.

[16] Goldstein, A. A., "Convex Programming in Hilbert Space," Bulletin of the American Mathematical Society 70 (1964) 709-710.

[17] Kyparisis, J. and A. V. Fiacco, "Second Order Sufficient Conditions for Strict and Isolated Local Minima in Nonlinear Programming," preprint, Florida International University, Department of Decision Sciences, Miami, Florida 1985.

[18] Levitin, E. S. and B. T. Polyak, "Constrained Minimization Problems," USSR Computational Mathematics and Mathematical Physics 6 (1966) 1-50.

[19] Luenberger, D. G., Linear and Nonlinear Programming, Second Edition, Addison Wesley Publishing Company, Reading, Massachusetts, 1984.

[20] McCormick, G. P., Nonlinear Programming Theory, Algorithms and Applications, John Wiley & Sons, New York, 1983.

[21] Pang, J. S. and D. Chan, "Iterative Methods for Variational and Complementarity Problems," Mathematical Programming 24 (1982) 284-313.

[22] Robinson, S. M., "Generalized Equations and Their Solutions, Part II: Applications to Nonlinear Programming," Mathematical Programming Study 19 (1982) 200-221.

[23] Tobin, R. L., "Sensitivity Analysis for Variational Inequalities," Journal of Optimization Theory and Applications 48 (1986) 191-204.

END
DATE
FILMED
6-88
DTIC